ELSEVIER Journal of Computational and Applied Mathematics 57 (1995) 57–76

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Favard theorem for reproducing kernels

Adhemar Bultheel[a,*], Pablo González-Vera[b], Erik Hendriksen[c], Olav Njåstad[d]

[a]*Department of Computer Science, Katholieke Universiteit Leuven, Celestijnen luan 200 A, B-3001 Leuven, Belgium*
[b]*Department of Mathematical Analysis, University of La Laguna, Tenerife, Spain*
[c]*Department of Mathematics, University of Amsterdam, Netherlands*
[d]*Department of Mathematics, University of Trondheim-NTH, Trondheim, Norway*

Received 23 October 1992; revised 10 March 1993

Abstract

Consider for $n = 0, 1, \ldots$ the nested spaces $\mathscr{L}_n$ of rational functions of degree n at most with given poles $1/\bar{\alpha}_i$, $|\alpha_i| < 1$, $i = 1, \ldots, n$. Let $\mathscr{L} = \bigcup_0^\infty \mathscr{L}_n$. Given a finite positive measure μ on the unit circle, we associate with it an inner product on $\mathscr{L}$ by $\langle f, g \rangle = \int f\bar{g}\,\mathrm{d}\mu$. Suppose $k_n(z, w)$ is the reproducing kernel for $\mathscr{L}_n$, i.e., $\langle f(z), k_n(z, w) \rangle = f(w)$, for all $f \in \mathscr{L}_n$, $|w| < 1$, then it is known that they satisfy a coupled recurrence relation.

In this paper we shall prove a Favard type theorem which says that if you have a sequence of kernel functions $k_n(z, w)$ which are generated by such a recurrence, then there will be a measure μ supported on the unit circle so that k_n is the reproducing kernel for $\mathscr{L}_n$. The measure is unique under certain extra conditions on the points α_i.

Keywords: Orthogonal rational functions; Favard theorem; Reproducing kernel

1. Introduction

We shall be concerned with nested spaces $\mathscr{L}_n$ for $n = 0, 1, \ldots$ which consist of rational functions spanned by a basis of partial Blaschke products $\{B_k\}_{k=0}^n$ where $B_0 = 1, B_n = B_{n-1}\zeta_n$ for $n = 1, 2, \ldots$ and the Blaschke factors ζ_n are defined by

$$\zeta_n(z) = \frac{\bar{\alpha}_n}{|\alpha_n|}\frac{\alpha_n - z}{1 - \bar{\alpha}_n z}, \quad |\alpha_n| < 1.$$

By convention, we set $\bar{\alpha}_n/|\alpha_n| = -1$ for $\alpha_n = 0$. Note that when $\alpha_k = 0$ for all k, then $B_n(z) = z^n$ and $\mathscr{L}_n$ is the space Π_n of polynomials of degree at most n. These spaces have been studied in connection with the Pick–Nevanlinna problem [21–24, 26–28] and in many applications [1–15, 17, 25, 30].

* Corresponding author.

SSDI 0377-0427(93)E0234-D

Consider next a finite positive measure μ (all measures in this paper will be finite and positive) on the unit circle $\boldsymbol{T} = \{z \in \mathbb{C}\colon |z| = 1\}$, normalized by $\int \mathrm{d}\mu = 1$, and define the inner product

$$\langle f, g\rangle_\mu = \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\overline{g(\mathrm{e}^{\mathrm{i}\theta})}\,\mathrm{d}\mu(\theta) = \int f(t)\overline{g(t)}\,\mathrm{d}\mu(t), \quad t = \mathrm{e}^{\mathrm{i}\theta} \in \boldsymbol{T}.$$

Let us denote an orthonormal system for $\mathscr{L}_n$ w.r.t. this inner product by $\{\phi_k\}_{k=0}^{n}$ with $\phi_0 \in \mathscr{L}_0$ and $\phi_k \in \mathscr{L}_k \backslash \mathscr{L}_{k-1}, k = 1, 2, \ldots, n$.

The kernel function

$$k_n(z, w) = \sum_{k=0}^{n} \phi_n(z)\overline{\phi_n(w)}$$

is reproducing in the sense that for any $f \in \mathscr{L}_n$ and for any $w \in \boldsymbol{D} = \{z \in \mathbb{C}\colon |z| < 1\}$

$$\langle f(t), k_n(t, w)\rangle_\mu = f(w).$$

It is well known [6] that the orthogonal functions ϕ_n satisfy some recurrence relation that generalizes the Szegő recurrence for polynomials orthogonal on the unit circle. In [7] we proved a Favard theorem for these ϕ_n. This means that if we are given a set of functions ϕ_n, generated by a recurrence relation of the type alluded to, then they are orthonormal with respect to a certain measure that can actually be constructed.

On the other hand, it is also known [4] that the kernels $k_n(z, w)$ satisfy a typical recurrence relation and in this paper we shall prove a Favard type theorem for the kernels, which says that if we are given a sequence of functions $\{k_j(z, w)\}_{j=0}^{n}$, $w \in \boldsymbol{D}$, which satisfy this particular type of recurrence relation, then they will be reproducing kernels for $\mathscr{L}_n$ with respect to some measure that will be constructed in the proof.

We treat the general case for arbitrary, not necessarily distinct, α_k in the unit disk $\boldsymbol{D}$. Note that if we choose all $\alpha_k = 0$, then $\mathscr{L}_n = \Pi_n$ are polynomial spaces. Also for the polynomial case this type of Favard theorem is new.

2. Definitions and notations

We shall consider several measures on the unit circle. For example, the normalized Lebesgue measure will be denoted by

$$\mathrm{d}\lambda(\theta) = \frac{\mathrm{d}\theta}{2\pi} = \mathrm{d}\lambda(t) = \frac{\mathrm{d}t}{2\pi \mathrm{i}t}, \quad t = \mathrm{e}^{\mathrm{i}t} \in \boldsymbol{T}.$$

The space $L_2(\mu)$ of square integrable functions (on $\boldsymbol{T}$, w.r.t. μ) will be denoted as L_2 instead of $L_2(\lambda)$ when the measure is the Lebesgue measure. The Hardy subspace of all L_2 functions with analytic extension to the open unit disc $\boldsymbol{D}$ is denoted by H_2. The other function classes L_p and H_p, $0 < p \leqslant \infty$ are also classical (see [16, 19, 20, 29]). In particular, the Nevanlinna class N is the set of ratios g/h with $g, h \in H_\infty$. This class N contains all H_p, $0 < p \leqslant \infty$.

The substar conjugate of a function is defined by

$$f_*(z) = \overline{f(1/\bar{z})}.$$

The (generalized) Poisson kernel is

$$P(z,w) = \frac{1-|w|^2}{(z-w)(z-w)_*}, \quad w \in \boldsymbol{D}.$$

Note that when $z \in \boldsymbol{T}$, this reduces to the usual definition

$$P(z,w) = \frac{1-|w|^2}{|z-w|^2}, \quad z \in \boldsymbol{T}, \quad w \in \boldsymbol{D}.$$

For $f_n \in \mathscr{L}_n$, we also define a superstar conjugate to mean

$$f_n^*(z) = B_n(z) f_{n*}(z) \in \mathscr{L}_n.$$

By $H(\boldsymbol{D})$ we mean the set of functions holomorphic in $\boldsymbol{D} \subset \mathbb{C}$.

The class of bounded analytic functions (Schur functions) is denoted by

$$\mathscr{B} = \{f \in H(\boldsymbol{D}): f(\boldsymbol{D}) \subset \boldsymbol{D}\}$$

and the class of positive real functions (Carathéodory functions) is denoted by

$$\mathscr{P} = \{f \in H(\boldsymbol{D}): \Re f(\boldsymbol{D}) > 0\}.$$

Recall that the Cayley transform $c(f) = (1-f)/(1+f)$ is a one-to-one map of $\mathscr{P}$ onto $\mathscr{B}$.

Let J be the 2×2 signature matrix $J = 1 \oplus -1$. A matrix $\theta = [\theta_{ij}] \in N^{2\times 2}$ is called J-unitary if $\theta_* J \theta = J$ a.e., where the substar for a matrix is defined by

$$\begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}_* = \begin{bmatrix} \theta_{11*} & \theta_{21*} \\ \theta_{12*} & \theta_{22*} \end{bmatrix}.$$

A matrix $\theta \in N^{2\times 2}$ is called J-contractive (in $\boldsymbol{D}$) if

$$\theta^{\mathrm{H}} J \theta \leqslant J, \text{ a.e. in } \boldsymbol{D}$$

where $^{\mathrm{H}}$ denotes the complex conjugate transpose and the inequality sign means that $J - \theta^{\mathrm{H}} J \theta$ is positive semi definite.

Following [17], we shall call matrices that are J-contractive in $\boldsymbol{D}$ and J-unitary on $\boldsymbol{T}$ simply J-inner matrices, since they naturally generalize the notion of a complex inner function. One can easily check that the class of J-inner functions is closed under multiplication. These matrices will play an essential role in this paper. We quote the following result from [12] to illustrate how very specific the properties of J-inner matrices are.

Theorem 2.1. *Let $\theta = [\theta_{ij}]$ be a J-inner matrix. Set $a = \theta_{11} - \theta_{12}$, $b = \theta_{11} + \theta_{12}$, $c = \theta_{22} - \theta_{21}$ and $d = \theta_{22} + \theta_{21}$. Then,*

(1) $|\det \theta| = 1$,

(2) $\theta^{-1} = J\theta_* J$,

(3) $\theta J \theta_* = J$,

(4) $\dfrac{1}{2}\left[\dfrac{a}{b} + \dfrac{a_*}{b_*}\right] = \dfrac{1}{bb_*} = \dfrac{1}{2}\left[\dfrac{c}{d} + \dfrac{c_*}{d_*}\right] = \dfrac{1}{dd_*}$,

(5) θ^{H} *is J-inner,*
(6) $b^{-1}, d^{-1} \in H_2$,
(7) $b^{-1}a, d^{-1}c \in \mathscr{P}$,
(8) $b^{-1}\, d$ *is inner.*

An example of a constant J-inner matrix is

$$\theta = \frac{1}{\sqrt{1-|\rho|^2}}\begin{bmatrix} 1 & \bar{\rho} \\ \rho & 1 \end{bmatrix}, \quad \rho \in \boldsymbol{D} \quad \text{and} \quad \theta^{-1} = \frac{1}{\sqrt{1-|\rho|^2}}\begin{bmatrix} 1 & -\bar{\rho} \\ -\rho & 1 \end{bmatrix}.$$

The Blaschke–Potapov factor

$$\begin{bmatrix} \zeta_n(z) & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{with } \zeta_n(z) \text{ a Blaschke factor}$$

is an example of a J-inner matrix of degree 1.

If $k_n(z,w)$ is the reproducing kernel for $\mathscr{L}_n$, then the normalized kernel $K_n(z,w)$ is defined as

$$K_n(z,w) = \frac{k_n(z,w)}{\sqrt{k_n(w,w)}} \Leftrightarrow k_n(z,w) = K_n(z,w)K_n(w,w)$$

(note that $k_n(w,w) = \sum_0^n |\phi_n(w)|^2 > 0$).

The kernels satisfy the following properties.

Property 2.2. *Let K_n be the normalized and k_n the nonnormalized reproducing kernel for $\mathscr{L}_n$, then (superstar for kernels is w.r.t. the first argument)*

(1) $k_n(w,w) > 0,\ K_n(w,w) > 0$,
(2) $k_n(w,z) = \overline{k_n(z,w)},\ K_n(w,z) = \overline{K_n(z,w)}$ *(sesqui-analytic),*
(3) $k_n(z,w) = B_n(z)\overline{B_n(w)}k_n(1/\bar{w}, 1/\bar{z})$, *i.e.,* $k_n^*(z,w) = k_n^*(w,z)$,
(4) $k_n(z,\alpha_n) = \overline{\phi_n^*(\alpha_n)}\phi_n^*(z)$.

Proof. Properties (1) and (2) are obvious from

$$k_n(z,w) = \sum_{k=0}^{n} \phi_k(z)\overline{\phi_k(w)}$$

while properties (3) and (4) were proved in [4]. □

For these kernels, the following recurrence has been derived [3, 4, 12].

Theorem 2.3. *Let $K_n(z,w)$ be the normalized (reproducing) kernel for $\mathscr{L}_n$. Then (superstar w.r.t. the first argument)*

$$\begin{bmatrix} K_n^*(z,w) \\ K_n(z,w) \end{bmatrix} = \theta_n(z,w)\begin{bmatrix} K_{n-1}^*(z,w) \\ K_{n-1}(z,w) \end{bmatrix}, \quad \begin{bmatrix} K_0^*(z,w) \\ K_0(z,w) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{2.1}$$

with

$$\theta_n(z,w) = c\begin{bmatrix} 1 & \bar{\rho}_n \\ \rho_n & 1 \end{bmatrix}\begin{bmatrix} \zeta_n(z) & 0 \\ 0 & 1 \end{bmatrix} d \begin{bmatrix} 1 & \bar{\gamma}_n \\ \gamma_n & 1 \end{bmatrix},$$

$$c = (1 - |\rho_n|^2)^{-1/2}, \qquad d = (1 - |\gamma_n|^2)^{-1/2},$$

$$\rho_n = \rho_n(w) = \overline{\phi_n(w)/\phi_n^*(w)} = \overline{K_n^*(w,\alpha_n)/K_n(w,\alpha_n)}$$

$$= \overline{K_n^*(\alpha_n, w)}/\overline{K_n(\alpha_n, w)},$$

$$\gamma_n = \gamma_n(w) = -\zeta_n(w)\rho_n(w).$$

The coefficients ρ_n and γ_n belong to $\boldsymbol{D}$ for $w \in \boldsymbol{D}$.

Proof. The proof can be found in the cited references where ρ_n was shown to be given by $\rho_n = \overline{\phi_n(w)/\phi_n^*(w)}$. This expression can be transformed with the properties given before

$$\rho_n(w) = \frac{\overline{\phi_n(w)}}{\overline{\phi_n^*(w)}} = \frac{\overline{k_n^*(w,\alpha_n)}}{\overline{k_n(w,\alpha_n)}} = \frac{\overline{k_n^*(\alpha_n,w)}}{k_n(\alpha_n,w)} = \frac{\overline{K_n^*(\alpha_n,w)}}{K_n(\alpha_n,w)},$$

which are the given expressions. □

If we introduce the kernels of the second kind $L_n(z,w)$ by

$$\begin{bmatrix} L_n^*(z,w) \\ -L_n(z,w) \end{bmatrix} = \theta_n(z,w)\begin{bmatrix} L_{n-1}^*(z,w) \\ -L_{n-1}(z,w) \end{bmatrix}, \quad \begin{bmatrix} L_0^*(z,w) \\ -L_0(z,w) \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

then clearly

$$\begin{bmatrix} K_n^* & L_n^* \\ K_n & -L_n \end{bmatrix} = \theta_n\theta_{n-1}\cdots\theta_1\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Thus, if we set $\Theta_n = \theta_n\theta_{n-1}\cdots\theta_1$, we get

$$\Theta_n = \frac{1}{2}\begin{bmatrix} K_n^* + L_n^* & K_n^* - L_n^* \\ K_n - L_n & K_n + L_n \end{bmatrix}. \tag{2.2}$$

The form of the recurrence relation implies the following property.

Corollary 2.4. *The normalized kernels K_n and L_n satisfy*

$$L_n(w,w) = K_n(w,w) = \prod_{k=1}^{n}\sqrt{\frac{1-|\gamma_k|^2}{1-|\rho_k|^2}}.$$

Proof. Using $\gamma_n = -\zeta_n\rho_n$, we can derive from the definition of $\theta_n(z,w)$ that

$$\theta_n(w,w) = \begin{bmatrix} (1-|\rho_n|^2)\zeta_n & \bar{\rho}_n(1-|\zeta_n|^2) \\ 0 & 1-|\gamma_n|^2 \end{bmatrix}\frac{1}{\sqrt{1-|\rho_n|^2}\sqrt{1-|\gamma_n|^2}}.$$

This implies that

$$\begin{bmatrix} K_n(w,w) \\ -L_n(w,w) \end{bmatrix} = \sqrt{\frac{1-|\gamma_n(w)|^2}{1-|\rho_n(w)|^2}} \begin{bmatrix} K_{n-1}(w,w) \\ -L_{n-1}(w,w) \end{bmatrix}$$

and because $K_0(w,w) = 1 = L_0(w,w)$, we get the expression that was claimed. □

The nonnormalized kernels satisfy a similar recurrence viz.

$$\begin{bmatrix} k_n^*(z,w) \\ k_n(z,w) \end{bmatrix} = t_n(z,w) \begin{bmatrix} k_{n-1}^*(z,w) \\ k_{n-1}(z,w) \end{bmatrix}, \quad \begin{bmatrix} k_0^*(z,w) \\ k_0(z,w) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{2.3}$$

with

$$t_n(z,w) = \sqrt{\frac{1-|\gamma_n(w)|^2}{1-|\rho_n(w)|^2}}\, \theta_n(z,w),$$

which follows easily from the previous corollary. As a consequence we also find that

$$k_n(w,w) = [K_n(w,w)]^2 = \prod_{k=1}^{n} \frac{1-|\gamma_k(w)|^2}{1-|\rho_k(w)|^2}.$$

Corollary 2.5. *If K_n is the normalized reproducing kernel for $\mathscr{L}_n$ w.r.t. some measure μ, then all the normalized reproducing kernels K_k for $k = 1, 2, \ldots, n-1$ are defined uniquely in terms of K_n.*

Proof. Because ρ_k is uniquely defined by K_k, we can invert the previous recurrence relation, which, by induction, is uniquely defined by K_n. □

Of course the same result holds also for the ordinary (nonnormalized) kernels. There is one more property about the reproducing kernels for $\mathscr{L}_n$ that we shall use later.

Property 2.6. *Given some measure, let $k_n(z,w)$ be the reproducing kernel for $\mathscr{L}_n$. Then there exists a sequence $\omega_j, j = 0, 1, \ldots$ in $\mathbf{D}$ such that the functions $k_j(z,\omega_j)$ $j = 0, 1, \ldots, n$ form a basis for $\mathscr{L}_n$, $n = 0, 1, \ldots$.*

Proof. Let us write the function $k_n(z,w)$ in terms of the basis of the finite Blaschke products

$$k_n(z,w) = a_0(w) + a_1(w)B_1(z) + \cdots + a_n(w)B_n(z).$$

The function $k_n(z,\omega_n)$ will be in $\mathscr{L}_n \backslash \mathscr{L}_{n-1}$ if $a_n(\omega_n) \neq 0$. Now clearly $\overline{a_n(w)} = k_n^*(\alpha_n, w)$ and because the latter is, considered as a function of w, an element from $\mathscr{L}_n$. It can therefore have at most n zeros. Thus it is always possible to select some ω_n such that $k_n^*(\alpha_n, \omega_n) \neq 0$, hence also $a_n(\omega_n) \neq 0$. □

3. Measures and interpolation

Let us define the kernel

$$D(z,w) = \frac{z+w}{z-w}.$$

Note the following relation with the Poisson kernel:

$$P(z,w) = \tfrac{1}{2}[D(z,w) + D(z,w)_*]$$

(substar w.r.t. the first argument), so that for $z \in \boldsymbol{T}$, $P(z,w) = \Re D(z,w)$.

With a measure μ on $\boldsymbol{T}$, we associate $\Omega \in \mathscr{P}$ by

$$\Omega(z) = \int D(t,z)\,\mathrm{d}\mu(t) + \mathrm{i}c, \quad c \in \mathbb{R}, \tag{3.1}$$

which belongs to H_p for all $p < 1$ [16, p. 34] and

$$\Re\Omega(z) = \int P(t,z)\,\mathrm{d}\mu(t)$$

has a nontangential limit to the unit circle a.e.,

$$\lim_{r\to 1^-} \Re\Omega(r\mathrm{e}^{\mathrm{i}\theta}) = \mu'(\mathrm{e}^{\mathrm{i}\theta}) \text{ a.e.,} \quad \mu'(\mathrm{e}^{\mathrm{i}\theta}) = \lim_{h\to 0} \frac{\mu((\theta-h,\theta+h))}{2h}.$$

Note that if $\int \mathrm{d}\mu = 1$, we get $\Omega(0) = 1 + \mathrm{i}c$. In fact, every $\Omega \in \mathscr{P}$ can be represented by an integral of this form, which is known as the Riesz–Herglotz representation. The relation between Ω and μ is one-to-one except for the real constant c, which is $c = \Im\Omega(0)$. Thus if $\int \mathrm{d}\mu = c_0 = 1$ and $c = 0$, then $\Omega(0) = 1$. In general, c can be chosen to make $\Omega(w) > 0$ for some $w \in \boldsymbol{D}$. With this particular choice of c (i.e., for $w = 0$ and $c = 0$), we shall denote the integral (3.1) by

$$\Omega(z) = \mathscr{T}_0(\mu), \quad \Omega(0) = 1 > 0.$$

Let $\mathscr{L}_n$ be defined by the set of points

$$A_n = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}.$$

Suppose we reorder them such that repeated points are brought together:

$$A_n = \{\underbrace{\beta_0, \ldots, \beta_0}_{\nu_0}, \underbrace{\beta_1, \ldots, \beta_1}_{\nu_1}, \ldots, \underbrace{\beta_m, \ldots, \beta_m}_{\nu_m}\}.$$

with $\beta_0 = 0$ and $\nu_0 \geqslant 0$, while $\nu_1, \ldots, \nu_m$ are all positive integers and $\sum_{k=0}^m \nu_k = n$. It is clear that a basis for $\mathscr{L}_n$ is given by

$$\{e_k\}_{k=0}^n = \{1, z, \ldots, z^{\nu_0}, (1-\bar\beta_1 z)^{-1}, \ldots, (1-\bar\beta_1 z)^{-\nu_1}, \ldots, (1-\bar\beta_m z)^{-1}, \ldots, (1-\bar\beta_m z)^{-\nu_m}\}.$$

Among other forms, a typical element in the Gram matrix for the latter basis has the form

$$\langle (1-\bar{\alpha}t)^{-k}, (1-\bar{\beta}t)^{-l}\rangle_\mu = \int \frac{1}{(1-\bar{\alpha}t)^k}\frac{t^l}{(t-\beta)^l}\,\mathrm{d}\mu(t). \tag{3.2}$$

One can easily check that

$$\frac{\mathrm{d}^k}{\mathrm{d}w^k}D(t,w) = 2(k!)t(t-w)^{-(k+1)}, \quad k \geqslant 1$$

and hence

$$\left[\frac{\mathrm{d}^k}{\mathrm{d}w^k}D(t,w)\right]_* = 2(k!)t^k(1-\bar{w}t)^{-(k+1)}, \quad k \geqslant 1$$

(substar w.r.t. t) so that one can derive that

$$\frac{\mathrm{d}^k}{\mathrm{d}w^k}\Omega(w) = \int \frac{\mathrm{d}^k}{\mathrm{d}w^k}D(t,w)\,\mathrm{d}\mu(t) = 2(k!)\int \frac{t}{(t-w)^{k+1}}\,\mathrm{d}\mu(t)$$

and

$$\overline{\frac{\mathrm{d}^k}{\mathrm{d}w^k}\Omega(w)} = \int \left[\frac{\mathrm{d}^k}{\mathrm{d}w^k}D(t,w)\right]_* \mathrm{d}\mu(t).$$

By partial fraction decomposition, one can see that integrals like (3.2), and hence the Gram matrix, will only depend upon values of

$$\left.\frac{\mathrm{d}^k}{\mathrm{d}w^k}\Omega(w)\right|_{w=\beta_s}$$

After checking all the details, one will have proved that the following is true.

Lemma 3.1. *Let μ and v be two measures on $\boldsymbol{T}$ and*

$$\Omega_\mu(z) = \mathscr{T}_0(\mu) \quad \text{and} \quad \Omega_v(z) = \mathscr{T}_0(v).$$

Then the inner product on $\mathscr{L}_n$ w.r.t. μ and w.r.t. v is the same if and only if Ω_μ interpolates Ω_v (in Hermite sense, taking repetition of points into account) in the point set $A_n^0 = \{0, \alpha_1, \ldots, \alpha_n\}$ which defines the space $\mathscr{L}_n$. Thus

$$\langle\cdot,\cdot\rangle_\mu = \langle\cdot,\cdot\rangle_v \text{ on } \mathscr{L}_n \iff \frac{\Omega_\mu(z) - \Omega_v(z)}{zB_n(z)} = g(z) \in H(\boldsymbol{D}).$$

The next lemma was proved in [12, p. 458].

Lemma 3.2. *Let μ be a measure on $\boldsymbol{T}$ and let $\Omega_\mu = \mathscr{T}_0(\mu)$. Define the positive real function*

$$\Omega_\mu^w(z) = \mathscr{T}_w(\mu) := \int \frac{D(t,z)}{P(t,w)}\,\mathrm{d}\mu(t) + \mathrm{i}c, \tag{3.3}$$

where P is the Poisson kernel and c is a real constant which normalizes Ω_μ^w, by $\Omega_\mu^w(w) > 0$. Then

$$\Omega_\mu^w(z) = \frac{\Omega_\mu(z)}{P(z,w)} + \frac{1}{1-|w|^2}\left(\frac{w}{z} - z\bar{w}\right)c_0, \quad c_0 = \int \mathrm{d}\mu.$$

Note 1. The choice of c which makes $\Omega_\mu^w(w) > 0$ is

$$\mathrm{i}c = \frac{1}{1-|w|^2}(\bar{w}c_1 - wc_{-1})$$

with $c_k = \int t^k \mathrm{d}\mu(t)$, the moments of μ.

Note 2. One can verify that $\Omega_\mu^w(w) = \Omega_\mu(0)$. Thus if $\int \mathrm{d}\mu = c_0 = 1 = \Omega_\mu(0)$, then $\Omega_\mu^w(w) = 1$ too.

Note 3. Taking the limit for $z \to 0$ the formula becomes

$$\Omega_\mu^w(0) = \frac{1}{1-|w|^2}[(1+|w|^2)c_0 - 2wc_{-1}].$$

The previous lemma has the following simple consequence which is a generalization of Lemma 3.1.

Corollary 3.3. *Let μ and ν be two measures on $\boldsymbol{T}$ and let*

$$\Omega_\mu^w(z) = \mathscr{T}_w(\mu) \quad \text{and} \quad \Omega_\nu^w(z) = \mathscr{T}_w(\nu)$$

with $\mathscr{T}_w$ defined as in the previous lemma. Then the inner product on $\mathscr{L}_n$ with respect to μ and with respect to ν is the same if and only if Ω_μ^w interpolates Ω_ν^w in the point set $A_n^w = \{w, \alpha_1, \ldots, \alpha_n\}$ in Hermite sense, i.e.,

$$\langle\cdot,\cdot\rangle_\mu = \langle\cdot,\cdot\rangle_\nu \text{ on } \mathscr{L}_n \Leftrightarrow \frac{\Omega_\mu^w(z) - \Omega_\nu^w(z)}{(z-w)B_n(z)} = g(z) \in H(\boldsymbol{D}).$$

Proof. With $\Omega_\mu = \mathscr{T}_0(\mu)$ and $\Omega_\nu = \mathscr{T}_0(\nu)$, we get from the previous lemma

$$\begin{aligned}\frac{\Omega_\mu^w(z) - \Omega_\nu^w(z)}{(z-w)B_n(z)} &= \frac{\Omega_\mu(z) - \Omega_\nu(z)}{(z-w)B_n(z)P(z,w)} \\ &= \frac{1-\bar{w}z}{1-|w|^2}\frac{\Omega_\mu(z) - \Omega_\nu(z)}{zB_n(z)},\end{aligned}$$

which will be in $H(\boldsymbol{D})$ if and only if Ω_μ interpolates Ω_ν in $A_n^0 = \{0, \alpha_1, \ldots, \alpha_n\}$. By Lemma 3.1, this is true if and only if $\langle\cdot,\cdot\rangle_\mu = \langle\cdot,\cdot\rangle_\nu$ on $\mathscr{L}_n$. □

4. The Pick–Nevanlinna algorithm

Like the Szegő polynomials are related to the Schur coefficient problem [10], the rational functions of this paper are related to the Pick–Nevanlinna interpolation problem. When the

Pick–Nevanlinna algorithm (sometimes called generalized Schur algorithm) is brought into a particular form, it will be clear from our derivation below that when we run it backward, we get in fact the recursion for reproducing kernels.

Let $S_0(z,w) \in \mathscr{B}$ be a given Schur function, which depends on some fixed parameter $w \in \boldsymbol{D}$ and for which $S_0(w,w) = 0$. We are also given a sequence of interpolation points $\alpha_1, \alpha_2, \ldots$, all in $\boldsymbol{D}$, and not necessarily distinct. For simplicity, we suppose that S_0 is not a finite Blaschke product with zeros $\alpha_1, \alpha_2, \ldots$ (otherwise the algorithm would end after a finite number of steps).

We describe the first step of the algorithm. It consists of a three stage transformation performed on S_0 to give $S_1 \in \mathscr{B}$:

$$S_1(z,w) = \tau_{31} \circ \tau_{21} \circ \tau_{11}(S_0(z,w)) = \tau_1(S_0(z,w)),$$

where

$$\tau_{11}: S_0 \mapsto S_1' = \frac{S_0 - \gamma_1}{1 - \bar{\gamma}_1 S_0}, \quad \gamma_1 = \gamma_1(w) = S_0(\alpha_1, w),$$

$$\tau_{21}: S_1' \mapsto S_1'' = S_1'/\zeta_1, \quad \zeta_1(z) = \frac{\bar{\alpha}_1}{|\alpha_1|} \frac{z - \alpha_1}{1 - \bar{\alpha}_1 z},$$

$$\tau_{31}: S_1'' \mapsto S_1 = \frac{S_1'' - \rho_1}{1 - \bar{\rho}_1 S_1''}, \quad \rho_1 = \rho_1(w) = S_1''(w,w).$$

S_1 will again be in $\mathscr{B}$ and it is zero for $z = w$. The first step is a bijection of $\mathscr{B}$ onto $\mathscr{B}$ which makes S_1' zero in $z = \alpha_1$. The second step divides out this zero in such a way that the result S_1'' is again in $\mathscr{B}$ and the third step "normalizes" S_1 so that it is zero in $z = w$. The last step is not really necessary, but we shall see later that it gives the recurrence we need.

The Pick–Nevanlinna algorithm now continues to do a similar transformation on S_1, using the interpolation point α_2 (which may be the same as α_1) etc.

$$\begin{aligned} S_k = \tau_k(S_{k-1}) &= \tau_{1k} \circ \tau_{2k} \circ \tau_{3k}(S_{k-1}) \\ &= \tau_{1k} \circ \tau_{2k}(S_k') \\ &= \tau_{1k}(S_k''), \quad k \geqslant 1. \end{aligned}$$

If S_0 is in $\mathscr{B}$ and not rational, then this will continue indefinitely and all ρ_k and γ_k will be in $\boldsymbol{D}$. One easily sees that

$$S_k'(w,w) = \frac{S_{k-1}(w,w) - \gamma_k}{1 - \bar{\gamma}_k S_{k-1}(w,w)} = -\gamma_k = \zeta_k(w) S_k''(w,w) = \zeta_k(w)\rho_k.$$

Conversely, given τ_k, $k = 1,2, \ldots, n$ we may choose $\Gamma_0 \in \mathscr{B}$ such that $\Gamma_0(w) = 0$ and generate

$$\Gamma_{k+1} = \tau_{n-k}^{-1}(\Gamma_k), \quad k = 0,1, \ldots, n-1.$$

The function Γ_n will be equal to S_0 if $\Gamma_0 = S_n$, but in general, when $\Gamma_0 \neq S_n$, Γ_n will still interpolate S_0 in the point set $A_n^w = \{w, \alpha_1, \ldots, \alpha_n\}$. More generally, Γ_{n-k} will be a partial solution to this interpolation problem since it will interpolate S_k in $\{w, \alpha_n, \alpha_{n-1}, \ldots, \alpha_{k+1}\}$.

As in [12], we now give a homogeneous form of the same algorithm. Let $S_0 = \Delta_{01}/\Delta_{02} \in \mathscr{B}$ with $\Delta_{01}, \Delta_{02} \in H(\boldsymbol{D})$ and Δ_{02} zerofree in $\boldsymbol{D}$. We place numerator and denominator in a row vector $\Delta_0 = [\Delta_{01}\ \Delta_{02}]$. Such a matrix function is called admissible. The set of admissible matrices is

$$\mathscr{A} = \{\Delta = [\Delta_1\ \Delta_2] : \Delta_1, \Delta_2 \in H(\boldsymbol{D}),\ \Delta_2(z) \neq 0 \text{ for } z \in \boldsymbol{D},\ \Delta_1/\Delta_2 \in \mathscr{B}\}.$$

Note that $\Delta \in \mathscr{A}$ implies $\Delta J \Delta^{\mathrm{H}} < 0$ ($J = 1 \oplus -1$) in $\boldsymbol{D}$. The transformation $S_{n-1} = \tau_n^{-1}(S_n)$ of the Pick–Nevanlinna algorithm can now be formulated as $\Delta_{n-1} = \Delta_n \theta_n$ ($S_k = \Delta_{k1}/\Delta_{k2}$) with

$$\theta_n(z,w) = c\begin{bmatrix} 1 & \bar\rho_n \\ \rho_n & 1 \end{bmatrix}\begin{bmatrix} \zeta_n(z) & 0 \\ 0 & 1 \end{bmatrix} d \begin{bmatrix} 1 & \bar\gamma_n \\ \gamma_n & 1 \end{bmatrix},$$

$$c = (1 - |\rho_n|^2)^{-1/2}, \qquad d = (1 - |\gamma_n|^2)^{-1/2},$$

$$\gamma_n = \gamma_n(w) = \Delta_{n-1,1}(\alpha_n, w)/\Delta_{n-1,2}(\alpha_n, w),$$

$$\rho_n = \rho_n(w) = -\lim_{z \to w} \gamma_n(z)/\zeta_n(z).$$

The normalization constants c and d are not necessary, but they turn θ_n into a J-inner matrix (if $w \in \boldsymbol{D}$), and thus also the product $\Theta_n(z,w) = \theta_n \cdots \theta_1$ will be J-inner. As we showed in Section 2, Θ_n can be written in the form (2.2) with $K_n(z,w)$ and $L_n(z,w)$ in $\mathscr{L}_n$ for fixed $w \in \boldsymbol{D}$. Now we use $\Theta_n^{-1} = J\Theta_{n*}J = B_{n*}J\Theta_n^* J$, and $B_{n*} = 1/B_n$ to get

$$\Theta_n^{-1} = \frac{1}{2}\begin{bmatrix} K_n + L_n & -K_n^* + L_n^* \\ -K_n + L_n & K_n^* + L_n^* \end{bmatrix} B_n^{-1}.$$

If we plug this into $\Delta_0 \Theta_n^{-1} = \Delta_n$, with $\Delta_0 = [1 - \Omega(z,w)\quad 1 + \Omega(z,w)]$, where $\Omega(z,w) = (1 - S_0(z,w))/(1 + S_0(z,w)) \in \mathscr{P}$ for $w \in \boldsymbol{D}$, or equivalently

$$S_0(z,w) = \frac{1 - \Omega(z,w)}{1 + \Omega(z,w)} \in \mathscr{B}, \quad \Omega(w,w) = 1,$$

then we get

$$\tfrac{1}{2}[1 - \Omega(z,w)\quad 1 + \Omega(z,w)]\begin{bmatrix} K_n + L_n & -K_n^* + L_n^* \\ -K_n + L_n & K_n^* + L_n^* \end{bmatrix} = B_n \Delta_n.$$

This implies that

$$[L_n - K_n\Omega\quad L_n^* + K_n^*\Omega] = B_n\Delta_n,$$

with first component

$$L_n(z,w) - K_n(z,w)\Omega(z,w) = B_n(z)\Delta_{n1}(z,w).$$

Now, since Θ_n is J-inner and since

$$K_n = (\Theta_n)_{21} + (\Theta_n)_{22} \quad \text{and} \quad L_n = (\Theta_n)_{22} - (\Theta_n)_{21},$$

we know by Theorem 2.1 that $1/K_n \in H_2$ and $L_n/K_n \in \mathscr{P}$ for $w \in \boldsymbol{D}$. Hence, setting $\Omega_n = L_n/K_n$, we get

$$\Omega_n(z,w) - \Omega(z,w) = B_n g_w(z)$$

with $g_w = \Delta_{n1}/K_n$. Because $\Delta_n \in H(\mathbf{D})$, and $\Delta_{n1}(w,w) = 0$, we may conclude that $g_w \in H(\mathbf{D})$ and $g_w(w) = 0$. This means that Ω_n interpolates Ω in the point set $A_n^w = \{w, \alpha_1, \ldots, \alpha_n\}$.

Note that when choosing $\Delta_n = [0 \quad 2]$, then $\Delta_0 = \Delta_n \Theta_n$ will give

$$\Omega_n(z,w) = \frac{L_n}{K_n} = \frac{\Delta_{02} - \Delta_{01}}{\Delta_{02} + \Delta_{01}}.$$

One can choose more generally any $\tilde{\Delta}_n \in \mathscr{A}$ with $\tilde{\Delta}_{n1}(w) = 0$ and generate $\tilde{\Delta}_0 = \tilde{\Delta}_n \Theta_n$ which will also give some

$$\tilde{\Omega}_n(z,w) = \frac{\tilde{\Delta}_{02} - \tilde{\Delta}_{01}}{\tilde{\Delta}_{02} + \tilde{\Delta}_{01}},$$

which will also interpolate Ω in A_n^w.

Thus instead of working with Schur functions S_0 and interpolating Schur functions Γ_n, we work with a positive real function $\Omega \in \mathscr{P}$ and find interpolating positive real functions $\Omega_n \in \mathscr{P}$.

With the Riesz–Herglotz representation theorem which relates positive real functions to positive measures on $\mathbf{T}$, we can derive from the previous results a statement about the approximation of measures.

Theorem 4.1. *Let μ be a measure on $\mathbf{T}$, normalized by $\int \mathrm{d}\mu = 1$ and define $\Omega = \Omega(z,w) = \mathscr{T}_w(\mu)(z)$ with $\mathscr{T}_w$ as in* (3.3). *Furthermore, define the absolutely continuous measure μ_n, depending on w by*

$$\mathrm{d}\mu_n(t,w) = \frac{P(t,w)\,\mathrm{d}\lambda(t)}{|K_n(t,w)|^2}, \tag{4.1}$$

where P is the Poisson kernel and K_n is obtained by the Pick–Nevanlinna algorithm applied to $\Delta_0 = [1 - \Omega \quad 1 + \Omega]$. Then on $\mathscr{L}_n$, the inner products $\langle \cdot,\cdot \rangle_\mu$ and $\langle \cdot,\cdot \rangle_{\mu_n}$ are the same. Consequently, if μ does not depend on w, then $\langle \cdot,\cdot \rangle_{\mu_n}$ will not depend on w for functions in $\mathscr{L}_n$.

Proof. We only have to show that

$$\Omega_n(z,w) = \frac{L_n(z,w)}{K_n(z,w)} = \mathscr{T}_w(\mu_n)$$

because, by construction with the Pick–Nevanlinna algorithm, we know that Ω_n interpolates Ω in $A_n^w = \{w, \alpha_1, \ldots, \alpha_n\}$. By Corollary 3.3 we then find equality of the inner products on $\mathscr{L}_n$.

Because the matrix Θ_n generated by the Pick–Nevanlinna algorithm is J-inner, we know that

$$\frac{1}{2}\left[\frac{L_n}{K_n} + \frac{L_{n*}}{K_{n*}}\right] = \frac{1}{K_n K_{n*}} = \frac{1}{2}[\Omega_n + \Omega_{n*}],$$

which for $z \in \mathbf{T}$ reduces to

$$\Re(\Omega_n(z,w)) = \Re\left(\frac{L_n(z,w)}{K_n(z,w)}\right) = \frac{1}{|K_n(z,w)|^2}.$$

We only have to check the normalization $\Omega_n(w,w) = 1 > 0$ (recall $\int d\mu = 1$). But exactly as in Corollary 2.4, we can derive that

$$L_n(w,w) = K_n(w,w) = \prod_{k=1}^{n} \sqrt{\frac{1-|\gamma_k|^2}{1-|\rho_k|^2}}.$$

The proof did not depend on K_n being normalized kernels but only on the structure of the J-inner matrix. Therefore, we may conclude that

$$\Omega_n(w,w) = \frac{L_n(w,w)}{K_n(w,w)} = 1.$$

Thus, the normalization required by $\mathscr{T}_w(\mu_n)$ is fulfilled and hence $\Omega_n = \mathscr{T}_w(\mu_n)$. This proves the theorem. □

It is not difficult to identify $K_n(z,w)$ now as the normalized reproducing kernel for $\mathscr{L}_n$ with respect to μ, and thus also with respect to μ_n.

Corollary 4.2. *With the notation of the previous theorem, it holds that $K_n(z,w)$ is the normalized reproducing kernel for the space $\mathscr{L}_n$ with respect to the measure μ, which is supposed not to depend on w, i.e., $k_n(z,w) \in \mathscr{L}_n$ as a function of z and*

$$\langle f(t), k_n(t,w)\rangle_\mu = f(w), \quad w \in \boldsymbol{D}, \quad f \in \mathscr{L}_n,$$

where $k_n(t,w) = K_n(w,w)\, K_n(t,w)$.

Proof. Note that $K_n(w,w) > 0$ and $1/K_n(t,w) \in H_2$, so that for any function $f \in \mathscr{L}_n$

$$\begin{aligned}
\langle f(t), k_n(t,w)\rangle_\mu &= \langle f(t), k_n(t,w)\rangle_{\mu_n} \\
&= \int f(t) \frac{K_{n*}(t,w) K_n(w,w) P(t,w)}{K_n(t,w) K_{n*}(t,w)} \, d\lambda(t) \\
&= \int \frac{f(t) K_n(w,w)}{K_n(t,w)} P(t,w) \, d\lambda(t) \\
&= f(w)
\end{aligned}$$

by the Poisson integral of an H_2 function.

Since for fixed $w \in \boldsymbol{D}$, the function $k_n(z,w) \in \mathscr{L}_n$ by construction, $k_n(z,w)$ is the reproducing kernel for $\mathscr{L}_n$. □

As a conclusion, we can say that, when the Pick–Nevanlinna algorithm is applied to

$$\Delta_0 = [1 - \mathscr{T}_w(\mu) \quad 1 + \mathscr{T}_w(\mu)],$$

then the resulting $\Theta_n(z,w)$ matrix will give us the normalized reproducing kernels $K_n(z,w)$ for $\mathscr{L}_n$ with respect to μ as well as the associated kernels $L_n(z,w)$.

Note that Θ_n depends only on the values of $\Omega(z) = \mathscr{T}_w(\mu)(z)$ for $z \in A_n^w$ (possibly using derivatives if points are repeated in A_n^w) where we have supposed that $\Omega(w) = 1$. Thus we obtain the same

Θ_n if we replace Ω by any other $\tilde{\Omega}$ which interpolates Ω in A_n^w. It will hold on $\mathscr{L}_n$ that $\langle\cdot,\cdot\rangle_\mu = \langle\cdot,\cdot\rangle_{\tilde{\mu}}$ where $\tilde{\Omega} = \mathscr{T}_w(\tilde{\mu})$. Such an arbitrary $\tilde{\Omega}$ can be written as

$$\tilde{\Omega} = \frac{\tilde{\Delta}_{02} - \tilde{\Delta}_{01}}{\tilde{\Delta}_{02} + \tilde{\Delta}_{01}},$$

where $\tilde{\Delta}_0 = \tilde{\Delta}_n \Theta_n$ with arbitrary $\tilde{\Delta}_n \in \mathscr{A}$, $\tilde{\Delta}_{n1}(w) = 0$. So we could choose $\mathrm{d}\tilde{\mu}(t) = P(t,w)\tilde{\mu}'(t)\,\mathrm{d}\lambda(t)$ with $\tilde{\mu}'(t)$ the nontangential limit to the unit circle of $\Re\tilde{\Omega}(z)$, or, when extended to the complex plane:

$$\tilde{\mu}'(z) = \tfrac{1}{2}[\tilde{\Omega}(z) + \tilde{\Omega}_*(z)],$$

since

$$\begin{aligned}\mathscr{T}_w(\tilde{\mu}) &= \int \frac{D(t,z)}{P(t,w)}\,\mathrm{d}\tilde{\mu}(t)\\ &= \frac{1}{2}\int D(t,z)[\tilde{\Omega}(t) + \tilde{\Omega}_*(t)]\,\mathrm{d}\lambda(t) = \tilde{\Omega}(z).\end{aligned}$$

5. Favard theorem

Now we shall try to reverse the process. Suppose, we are given some numbers $\alpha_n, n = 1,2, \ldots$ all in $\boldsymbol{D}$ and some numbers ρ_n which are also in $\boldsymbol{D}$ for $n = 1,2, \ldots$. These ρ_n may depend upon the complex parameter w. The dependence is for the moment unspecified. Then we generate some functions $k_n(z,w)$ by a recurrence relation that is formally the same as the recurrence relation for the reproducing kernels. As a function of z, these functions $k_n(z,w)$ will be in $\mathscr{L}_n$ by construction. What can we say about these functions without further specifying what the dependence is on w? Eventually, we shall of course want them to be reproducing kernels for $\mathscr{L}_n$ with respect to some measure. We shall start however with some simple lemmas where the dependence upon w is irrelevant.

Lemma 5.1. *Suppose $k_n(z,w)$ are functions in $\mathscr{L}_n$ depending on some parameter $w \in \boldsymbol{D}$ satisfying the recurrence relation*

$$\begin{aligned}&k_0(z,w) = 1,\\ &k_n(z,w) = e_n(w)[\lambda_n(z,w)k_{n-1}^*(z,w) + \hat{\lambda}_n(z,w)k_{n-1}(z,w)], \quad n = 1,2, \ldots ,\\ &\gamma_n(w) = -\zeta_n(w)\rho_n(w), \quad \rho_n(w) \in \boldsymbol{D},\\ &e_n(w) = (1 - |\rho_n(w)|^2)^{-1},\\ &\lambda_n(z,w) = \rho_n(w)\zeta_n(z) + \gamma_n(w) = \rho_n(w)[\zeta_n(z) - \zeta_n(w)] \in \mathscr{L}_1,\\ &\hat{\lambda}_n(z,w) = \rho_n(w)\zeta_n(z)\overline{\gamma_n(w)} + 1 = 1 - |\rho_n(w)|^2\zeta_n(z)\overline{\zeta_n(w)} \in \mathscr{L}_1.\end{aligned}$$

Then $k_n^*(z,w)$ *satisfies (the superstar is with respect to the first argument)*

$$k_0^*(z,w) = 1,$$
$$k_n^*(z,w) = e_n(w)[\sigma_n(z,w)k_{n-1}^*(z,w) + \hat{\sigma}_n(z,w)k_{n-1}(z,w)], \quad n = 1,2, \ldots,$$
$$\sigma_n(z,w) = \zeta_n(z) + \gamma_n(w)\overline{\rho_n(w)} = \zeta_n(z)\zeta_n(w)|\rho_n(w)|^2 = \hat{\lambda}_n^*(z,w),$$
$$\hat{\sigma}_n(z,w) = \zeta_n(z)\overline{\gamma_n(w)} + \overline{\rho_n(w)} = \overline{\rho_n(w)}[1 - \zeta_n(z)\overline{\zeta_n(w)}] = \lambda_n^*(z,w).$$

Proof. The formulation of the lemma is so explicit that its proof is trivial. □

Note that the previous result can be reformulated as

$$\begin{bmatrix} k_n^*(z,w) \\ k_n(z,w) \end{bmatrix} = t_n(z,w)\begin{bmatrix} k_{n-1}^*(z,w) \\ k_{n-1}(z,w) \end{bmatrix}, \quad \begin{bmatrix} k_0^*(z,w) \\ k_0(z,w) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

with $t_n(z,w)$ of exactly the same form as the matrix t_n of (2.3). This implies that Corollary 2.4 is applicable here in its reformulation for the functions k_n. Thus the k_n as generated above satisfy

$$k_n(w,w) = \prod_{k=1}^{n} \frac{1 - |\gamma_k(w)|^2}{1 - |\rho_k(w)|^2} > 0.$$

Hence we may consider the normalized versions $K_n(z,w) = k_n(z,w)/\sqrt{k_n(w,w)}$ and these satisfy the recurrence

$$\begin{bmatrix} K_n^*(z,w) \\ K_n(z,w) \end{bmatrix} = \theta_n(z,w)\begin{bmatrix} K_{n-1}^*(z,w) \\ K_{n-1}(z,w) \end{bmatrix}, \quad \begin{bmatrix} K_0^*(z,w) \\ K_0(z,w) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where $\theta_n(z,w)$ has the same form as in Section 2.

Lemma 5.2. *With the previous notation* $1/K_n(z,w) \in H_2$ *and hence also* $1/k_n(z,w) \in H_2$.

Proof. This follows for K_n directly from the θ_n being J-inner and the properties of Theorem 2.1. Together with the previous lemma this implies that the property also holds for k_n. □

Lemma 5.3. *Let* k_n *be generated as in the previous lemmas. Then*

$$\overline{\rho_n(w)} = \frac{k_n^*(\alpha_n,w)}{k_n(\alpha_n,w)} = \frac{K_n^*(\alpha_n,w)}{K_n(\alpha_n,w)}.$$

Proof. Note that

$$k_n^*(\alpha_n,w) = e_n(w)[\sigma_n(\alpha_n,w)k_{n-1}^*(\alpha_n,w) + \hat{\sigma}_n(\alpha_n,w)k_{n-1}(\alpha_n,w)],$$
$$\sigma_n(\alpha_n,w) = -\zeta_n(w)|\rho_n(w)|^2, \quad \hat{\sigma}_n(\alpha_n,w) = \overline{\rho_n(w)},$$
$$k_n(\alpha_n,w) = e_n(w)[\lambda_n(\alpha_n,w)k_{n-1}^*(\alpha_n,w) + \hat{\lambda}_n(\alpha_n,w)k_{n-1}(\alpha_n,w)],$$
$$\lambda_n(\alpha_n,w) = -\rho_n(w)\zeta_n(w), \quad \hat{\lambda}_n(\alpha_n,w) = 1.$$

Taking the ratio $k_n^*(\alpha_n,w)/k_n(\alpha_n,w)$ gives precisely $\overline{\rho_n(w)}$. □

As a consequence of this, we can, as in the case of reproducing kernels conclude that k_n will completely define all the previous k_j for $j = n-1, n-2, \dots, 0$ and similarly K_n will define all the previous ones. Thus if k_n is reproducing kernel for $\mathscr{L}_n$ with respect to some measure, then k_j will be reproducing kernels for $\mathscr{L}_n$, $j = n-1, n-2, \dots, 1$ with respect to the same measure.

This is about as far as we can get without further specification of how ρ_n depends upon w. For an arbitrary sequence of numbers $\rho_k(w)$, depending on w and satisfying $|\rho_k(w)| < 1$, one may not expect that the corresponding Θ_n matrix contains (normalized) reproducing kernels for $\mathscr{L}_n$ with respect to any measure whatsoever.

For arbitrary $\rho_k(w)$, $k = 1, \dots, n$, one can build Θ_n from which we can extract $K_n = (\Theta_n)_{21} + (\Theta_n)_{22}$ and the corresponding measure μ_n as in Theorem 4.1. We then do have that

$$\langle f(t), k_n(t,w)\rangle_{\mu_n} = f(w) \quad \forall f \in \mathscr{L}_n, \tag{5.1}$$

where $k_n(z,w) = K_n(w,w)K_n(z,w)$. However, this μ_n will depend on w and therefore we cannot conclude from (5.1) that k_n is a reproducing kernel since although it reproduces any $f \in \mathscr{L}_n$, it does so only in the special point w on which μ_n depends.

If $\phi_0, \dots, \phi_n$ is an orthonormal basis for $\mathscr{L}_n$, then the kernels are

$$k_n(z,w) = \sum_{k=0}^{n} \phi_k(z)\overline{\phi_k(w)}$$

and this reflects a specific symmetry in z and w. It implies for example that as a function of w, $\overline{k_n(z,w)}$ should be in $\mathscr{L}_n$. In general, a reproducing kernel should be sesqui-analytic, that is $\overline{k_n(z,w)} = k_n(w,z)$ and more specifically, in $\mathscr{L}_n$ all the relations given in Property 2.2 should hold. This means that the way in which $k_n(z,w)$ depends upon w is very special, and one should not expect that the choice of arbitrary $\rho_k(w)$, which depend in some exotic way on w, will provide this. One an easily check this by considering the simple case of $n = 1$ for example.

So we shall have to introduce the notion of a sequence $\rho_k(w)$ having the property that the corresponding k_n are indeed reproducing kernels. We shall say that such a sequence $\rho_k(w)$ has the reproducing kernel (RK) property.

Since the $k_n(z,w)$ as they were generated in the previous lemmas depend upon w via $\rho_i(w)$ in a very complex way, it is not easy to find conditions on how the coefficients $\rho_i(w)$ should depend upon w to ensure that $\overline{k_n(z,w)}$, as a function of w, is in $\mathscr{L}_n$. The reader is invited to try and check this for the simplest possible case $n = 1$.

It is yet an open problem to find a direct and simple characterisation of the $\rho_i(w)$ having the RK property. For the moment we content ourselves with a characterisation that is in the line of this paper and shall formulate some equivalent conditions. Unfortunately, none of these will give a direct characterization of how the coefficients ρ_k should depend on w. If such a characterization exists, it is still to be found.

As explained in the previous lemmas, there is a one-to-one correspondence between the coefficients $\{\rho_i(w)\colon i, \dots, n\}$, the functions $\{k_i(z,w)\colon i = 1, \dots, n\}$, the normalized functions $\{K_i(z,w)\colon i = 1, \dots, n\}$ and the J-inner matrices $\{\Theta_i(z,w)\colon i = 1, \dots, n\}$. We shall say that one of these (and therefore also all the others) has the RK property if on $\mathscr{L}_n$, the inner product $\langle \cdot, \cdot \rangle_{\mu_n}$ is independent of w, where μ_n is the measure defined in terms of the $K_n(z,w)$ by an expression like (4.1).

It is an immediate consequence of Theorem 4.1 that the $\rho_i(w)$ will have the RK property if they can be generated by the Pick–Nevanlinna algorithm applied to some $\Delta_0 \in \mathscr{A}$, $\Delta_{01}(w) = 0$ which is of the form

$$\Delta_0 = [1 - \mathscr{T}_w(\tilde{\mu}) \quad 1 + \mathscr{T}_w(\tilde{\mu})], \tag{5.2}$$

where

$$\mathscr{T}_w(\tilde{\mu})(z) = \frac{\Omega(z)}{P(z,w)} + \frac{1}{1-|w|^2}\left(\frac{w}{z} - z\bar{w}\right), \quad \Omega = \mathscr{T}_0(\tilde{\mu}) \tag{5.3}$$

for some measure $\tilde{\mu}$ satisfying $\int \mathrm{d}\tilde{\mu} = 1$ and independent of w. Since $L_n(z,w)/K_n(z,w)$ as generated by the Pick–Nevanlinna algorithm shall interpolate this $\mathscr{T}_w(\tilde{\mu})$ in A_n^w and thus also $L_n(z,0)/K_n(z,0)$ will interpolate $\Omega = \mathscr{T}_0(\tilde{\mu})$ in A_n^0, we see that $\Theta_n(z,w)$ shall have the RK property if

$$\frac{L_n(z,w)}{K_n(z,w)} = \frac{L_n(z,0)/K_n(z,0)}{P(z,w)} + \frac{1}{1-|w|^2}\left(\frac{w}{z} - z\bar{w}\right) \quad \text{for } z \in A_n^w.$$

In view of the comments given before Theorem 4.1, the Θ_i will also have the RK property if there exists some $\tilde{\Delta}_n \in \mathscr{A}$ with $\tilde{\Delta}_{n1}(w) = 0$ and $\tilde{\Delta}_0 = \tilde{\Delta}_n\Theta_n$ of the form (5.2) and (5.3).

We can use now $\tilde{\Delta}_n(z,w) = [\tilde{S}_n(z,w) \quad 1]$ with $\tilde{S}_n(z,w) \in \mathscr{B}$ and $\tilde{S}_n(w,w) = 0$, to get

$$\begin{aligned} \tilde{\Delta}_0 &= \tilde{\Delta}_n\Theta_n \\ &= \tfrac{1}{2}[\tilde{S}_n \quad 1]\begin{bmatrix} K_n^* + L_n^* & K_n^* - L_n^* \\ K_n - L_n & K_n + L_n \end{bmatrix} \\ &= \tfrac{1}{2}[\tilde{S}_n(K_n^* + L_n^*) + (K_n - L_n) \quad \tilde{S}_n(K_n^* - L_n^*) + (K_n + L_n)]. \end{aligned}$$

If this has to be of the form (5.2), then

$$\mathscr{T}_w(\tilde{\mu}) = \frac{\tilde{\Delta}_{02} - \tilde{\Delta}_{01}}{\tilde{\Delta}_{02} + \tilde{\Delta}_{01}} = \frac{L_n - \tilde{S}_n L_n^*}{K_n + \tilde{S}_n K_n^*}.$$

We may thus conclude that Θ_i, $i = 1, \ldots, n$ will have the RK property if there exists some function $\tilde{S}_n(z,w) \in \mathscr{B}$, which may depend upon a parameter w and which satisfies $\tilde{S}_n(w,w) = 0$, such that the function $\Omega_n(z)$, defined by

$$\tilde{\Omega}_n(z) = \left[\frac{L_n(z,w) - \tilde{S}_n(z,w)L_n^*(z,w)}{K_n(z,w) - \tilde{S}_n(z,w)K_n^*(z,w)} - \frac{z^{-1}w - z\bar{w}}{1-|w|^2}\right]P(z,w)$$

belongs to $\mathscr{P}$ and is independent of w.

We now have a Favard type theorem.

Theorem 5.4. *Let the $k_n(z,w)$ be generated as in the previous lemmas and let $K_n(z,w) = k_n(z,w)/\sqrt{k_n(w,w)}$ be their normalized versions. Suppose the $\rho_n(w)$ form a sequence with the RK property. Then there exists a Borel measure on $\boldsymbol{T}$ such that for $n = 0, 1, 2, \ldots$ the function $k_n(z,w)$ is a reproducing kernel for $\mathscr{L}_n$. Thus there is a measure μ such that for $n = 0, 1, 2, \ldots$*

$$\langle f(z), k_n(z,w)\rangle_\mu = f(w) \quad \forall f \in \mathscr{L}_n, \quad \forall w \in \boldsymbol{D}.$$

ɿtional functions $\bigcup_{n=0}^{\infty}\mathscr{R}_n$ where $\mathscr{R}_n = \mathscr{L}_n + \mathscr{L}_{n*}$ and $\mathscr{L}_{n*} = \{f_*: f \in \mathscr{L}_n\}$, are dense in the ˉ) of continuous functions on $\boldsymbol{T}$, then the measure μ is unique.

. the ρ_n have the RK property, then $\mu_n^w(t) = \mu_n(t,w)$ as defined in (4.1) will define an inner .ct $\langle\cdot,\cdot\rangle_{\mu_n}$ which on $\mathscr{L}_n$ will be independent of w, which implies as in Corollary 4.2 that the .,w) is a reproducing kernel for $\mathscr{L}_n$ with respect to $\mu_n(t) = \mu_n^0(t) = \mu_n(t,0)$. Because by the revious lemma, the kernel k_n defines all the previous ones, we shall also have that $k_j(z,w)$ is reproducing kernel for $\mathscr{L}_j$ with respect to the measure $\mu_n(t)$ for $j = n-1, n-2, \ldots$.

We can now use the same reasoning as in the case of the Favard theorem for the orthogonal functions [7] or for orthogonal polynomials [18]. Since the distribution functions

$$\mu_n(t) = \int_0^t \frac{P(\mathrm{e}^{\mathrm{i}\theta},0)}{|K_n(\mathrm{e}^{\mathrm{i}\theta},0)|^2}\,\mathrm{d}\lambda(\theta) = \int_0^t \frac{\mathrm{d}\lambda(\theta)}{|K_n(\mathrm{e}^{\mathrm{i}\theta},0)|^2}$$

are increasing functions and uniformly bounded ($\int \mathrm{d}\mu_n = 1$, because $\mathscr{T}_0(\mu_n) = \Omega_n(z) = L_n(z,0)/K_n(z,0)$ and $\Omega_n(0) = 1$ and $\int \mathrm{d}\mu = c_0 = \Omega_n(0)$), there exists a subsequence such that

$$\lim_{k\to\infty} \mu_{n_k}(\theta) = \mu(\theta) \quad \text{and} \quad \lim_{k\to\infty} \int f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\mu_{n_k}(\theta) = \int f\,\mathrm{d}\mu$$

for all f continuous on $\boldsymbol{T}$. Thus, for $n = 0, 1, \ldots$, the kernels $k_n(z,w)$ are all reproducing in $\mathscr{L}_n$ with respect to this measure μ.

To prove the uniqueness, we note that, because these k_n are reproducing kernels,we may apply Property 2.6. Thus there exists a sequence of complex numbers $\omega_n, n = 0, 1, \ldots$ such that the sequence of functions $k_n(z,\omega_n)$, $n = 0, 1, \ldots$ forms a basis for $\mathscr{L}_\infty$. Thus we may define a linear bounded functional Φ on $\mathscr{R}_\infty$ (hence, because of the denseness also in $C(\boldsymbol{T})$) by means of

$$\Phi(k_i(z,\omega_i)k_{j*}(z,\omega_j)) = \int k_i(z,\omega_i)\overline{k_j(z,\omega_j)}\,\mathrm{d}\mu = k_m(\omega_j,\omega_i),$$

where $m = \min\{i,j\}$. By the Riesz representation theorem of bounded linear functionals, it follows that μ is unique. □

Note. As in the Favard theorem for the orthogonal functions [7], the rationals being dense in $C(\boldsymbol{T})$ is only a sufficient condition for the uniqueness of the measure. The denseness of the rationals in $C(\boldsymbol{T})$ is equivalent with the Blaschke condition $\sum(1-|\alpha_k|) = \infty$. In the polynomial case where all $\alpha_k = 0$ and $\mathscr{L}_n = \Pi_n$, this condition is always satisfied. It is well known that the trigonometric polynomials are dense in $C(\boldsymbol{T})$. Also in the case where there is only a finite number of different α_k which are repeated cyclically, this condition is satisfied.

References

[1] N.I. Akhiezer, *The Classical Moment Problem* (Oliver and Boyd, Edinburgh, 1969; originally published: Moscow, 1961).

[2] A. Bultheel, On a special Laurent–Hermite interpolation problem, in: L. Collatz, G. Meinardus and H. Werner, Eds, *Numerische Methoden der Approximazionstheorie 6*, Internat. Ser. Numer. Math. **59** (Birkhäuser, Basel, 1981) 63–79.

[3] A. Bultheel and P. Dewilde, Orthogonal functions related to the Nevanlinna–Pick problem, in: P. Dewilde, Ed. *Proc. 4th Internat. Conf. on Math. Theory of Networks and Systems, Delft* (Western Periodicals, North-Hollywood, 1979) 207–212.

[4] A. Bultheel, P. González-Vera, E. Hendriksen and O. Njåstad, A Szegő theory for rational functions, Technical Report TW131, Department of Computer Science, K.U. Leuven, 1990.

[5] A. Bultheel, P. González-Vera, E. Hendriksen and O. Njåstad, Orthogonal rational functions similar to Szegő polynomials, in: C. Brezinski, L. Gori and A. Ronveaux, Eds., *Orthogonal Polynomials and Their Applications*, IMACS Ann. Comput. Appl. Math. **9** (Baltzer, Basel, 1991) 195–204.

[6] A. Bultheel, P. González-Vera, E. Hendriksen and O. Njåstad, The computation of orthogonal rational functions and their interpolating properties, *Numer. Algorithms* **2**(1) (1992) 85–118.

[7] A. Bultheel, P. González-Vera, E. Hendriksen and O. Njåstad, A Favard theorem for orthogonal rational functions on the unit circle, *Numer. Algorithms* **3** (1992) 81–89.

[8] A. Bultheel, P. González-Vera, E. Hendriksen and O. Njåstad, A moment problem associated to rational Szegő functions, *Numer. Algorithms* **3** (1992) 91–104.

[9] A. Bultheel, P. González-Vera, E. Hendriksen and O. Njåstad, Orthogonal rational functions and quadrature on the unit circle, *Numer. Algorithms* **3** (1992) 105–116.

[10] A. Bultheel, P. González-Vera, E. Hendriksen and O. Njåstad, Moment problems and orthogonal functions, *J. Comput. Appl. Math.* **48** (1–2) (1993) 49–68.

[11] P. Delsarte, Y. Genin and Y. Kamp, On the role of the Nevanlinna–Pick problem in circuit and system theory, *Internat. J. Circuit Theory Appl.* **9** (1981) 177–187.

[12] P. Dewilde and H. Dym, Schur recursions, error formulas, and convergence of rational estimators for stationary stochastic sequences, *IEEE Trans. Inform. Theory* **IT-27** (1981) 446–461.

[13] P. Dewilde and H. Dym, Lossless inverse scattering, digital filters, and estimation theory, *IEEE Trans. Inform. Theory* **IT-30** (1984) 644–662.

[14] P. Dewilde, A. Viera and T. Kailath, On a generalized Szegő–Levinson realization algorithm for optimal linear predictors based on a network synthesis approach, *IEEE Trans. Circuit Systems* **25** (1978) 663–675.

[15] M.M. Djrbashian, A survey on the theory of orthogonal systems and some open problems, in: P. Nevai, Ed., *Orthogonal Polynomials: Theory and Practice*, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **294** (Kluwer, Boston, MA, 1990) 135–146.

[16] P.L. Duren, *The Theory of H^p Spaces*, Pure Appl. Math. **38** (Academic Press, New York, 1970).

[17] H. Dym, *J Contractive Matrix functions, Reproducing Kernel Hilbert Spaces and Interpolation*, CBMS Regional Conf. Ser. in Math. **71** (Amer. Mathematical Soc., Providence, RI, 1989)

[18] T. Erdélyi, P. Nevai, J. Zhang and J.S. Geronimo, A simple proof of "Favard's theorem" on the unit circle, *Atti Sem. Mat. Fis. Univ. Modena* **29** (1991) 41–46.

[19] K. Hoffman, *Banach Spaces of Analytic Functions* (Prentice-Hall, Englewood Cliffs, NJ, 1962).

[20] P. Koosis, *Introduction to H^p Spaces*, London Math. Soc. Lecture Notes **40** (Cambridge Univ. Press, Cambridge, 1980).

[21] R. Nevanlinna, Über beschränkte Funktionen die in gegebenen Punkten vorgeschriebene Werte annehmen, *Ann. Acad. Sci. Fenn. Ser. A.* **13**(1) (1919).

[22] R. Nevanlinna, Asymptotische Entwickelungen beschränkter Funktionen und das Stieltjessche Momentenproblem, *Ann. Acad. Sci. Fenn. Ser. A.* **18**(5) (1922).

[23] R. Nevanlinna, Kriterien für die Randwerte beschränkter Funktionen, *Math. Z.* **13** (1922) 1–9.

[24] R. Nevanlinna, Über beschränkte analytische Funktionen, *Ann. Acad. Sci. Fenn. Ser. A.* **32**(7) (1929).

[25] O. Njåstad, Multipoint Padé approximation and orthogonal rational functions, in: A. Cuyt, Ed., *Nonlinear Numerical Methods and Rational Approximation* (Reidel, Dordrecht, 1988) 258–270.

[26] G. Pick, Über die Beschränkungen analytischen Funktionen welche durch vorgegebene Funktionswerte bewirkt werden, *Math. Ann.* **77** (1916) 7–23.

[27] G. Pick, Über die Beschränkungen analytischen Funktionen durch vorgegebene Funktionswerte, *Math. Ann.* **78** (1918) 270–275.
[28] G. Pick, Über beschränkte Funktionen mit vorgeschriebenen Wertzuordnungen, *Ann. Acad. Sci. Fenn. Ser. A* **15**(3) (1920).
[29] W. Rudin, *Real and Complex Analysis* (McGraw-Hill, New York, 2nd ed., 1974).
[30] J.L. Walsh, *Interpolation and Approximation*, AMS Colloq. Publ. **20** (Amer. Mathematical Soc., Providence, RI, 3rd ed., 1960; 1st ed., 1935).